B16

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 608 004 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **94103625.3**

(22) Date of filing: **11.08.87**

(51) Int. Cl.5: **C12Q 1/68**, //G01N33/574, G01N33/68,C07H21/00

This application was filed on 09 - 03 - 1994 as a divisional application to the application mentioned under INID code 60.

(30) Priority: **11.08.86 US 895163**

(43) Date of publication of application: **27.07.94 Bulletin 94/30**

(60) Publication number of the earlier application in accordance with Art.76 EPC: **0 259 031**

(84) Designated Contracting States: **AT BE CH DE ES FR GB GR IT LI LU NL SE**

(71) Applicant: **MASSACHUSETTS EYE & EAR INFIRMARY**
243 Charles Street
Boston, MA 02114(US)
Applicant: **WHITEHEAD INSTITUTE**
Nine Cambridge Center
Cambridge, MA 02130(US)

(72) Inventor: **Dryja, Thaddeus P.**
85 Forbes Road
Milton, Massachusetts 02186(US)
Inventor: **Friend, Stephen**
14 Spencer Avenue
Somerville, Massachusetts 02143(US)

(74) Representative: **Wright, Simon Mark et al**
Kilburn & Strode
30 John Street
London WC1N 2DD (GB)

(54) Human DNA in the diagnosis of retinoblastoma.

(57) Genetic material corresponding to a normal human retinoblastoma is compared with DNA from a patient to diagnose the presence of defective retinoblastoma alleles.

EP 0 608 004 A2

This invention relates to methods of detection and treatment of a defective human gene related to cancer, in particular, retinoblastoma.

In M. Lalande *et al*, Cancer Genet Cytogenet 23: 151-157 (1986) the cloning of different DNA segments in human chromosomes has lead to a set of probes that might be useful in the diagnosis of human retinoblastoma.

Retinoblastoma is a neoplastic condition of the retinal cells, observed almost exclusively in children between the ages of 0 and 4 years. If untreated, the malignant neoplastic retinal cells in the intraocular tumor travel to other parts of the body, forming foci of uncontrolled growth which are always fatal. The current treatment for a retinoblastoma is enucleation of the affected eye if the intraocular tumor is large; for small intraocular tumors, radiation therapy, laser therapy, or cryotherapy is preferred. There is no known successful treatment for metastatic retinoblastoma. Hence, early diagnosis of retinoblastoma to allow treatment before the tumor spreads outside the eye is crucial.

There is evidence that retinoblastoma is caused by the functional loss of both homologous copies of the retinoblastoma (Rb) gene. Thus, individuals carrying one defective allele of the Rb gene are genetically predisposed to the disease. Children who have had one eye affected by retinoblastoma or who are related to someone with retinoblastoma may be genetically predisposed and therefore at risk of developing the disease. These individuals routinely are tested for retinoblastoma every 2-3 months by an ocular examination procedure which requires placing the child under general anesthesia.

The invention relates to the use of genetic material corresponding to a normal human retinoblastoma gene or a unique subregion thereof in the preparation of material for use in a method of screening human patients comprising comparing the DNA of said patients with the said gene or subregion. Also encompassed are vectors comprising genetic material corresponding to a normal retinoblastoma gene, or a unique subregion thereof.

In general, the invention features a method of screening human patients by comparing the DNA of these patients with the isolated normal human retinoblastoma (Rb) gene or a unique subregion thereof (the term "unique subregion" means a DNA sequence found in the Rb gene and not elsewhere in the human genome). This comparison allows detection of defective Rb alleles in the patients, to determine whether these patients need continual monitoring by the conventional examination procedure. More importantly, this comparison will identify those patients who do not have a defective Rb allele and thus are not at risk of developing retinoblastoma and do not have to be examined by the conventional procedure.

Preferably, the comparison between the patient's DNA and the normal Rb gene involves testing the patient's DNA with the isolated Rb gene to detect either large deletions or, alternatively, small deletions or point mutations in the Rb locus. To test for large deletions in a patient's Rb allele, the patient's DNA preferably is analyzed by DNA hybridization using probes made from the isolated normal Rb gene. According to the invention, small deletions or point mutations preferably are detected by either of two techniques. The nucleotide sequences of the patients' Rb alleles and the normal Rb gene can be determined and compared for differences. Alternatively, the patient's DNA is probed with the normal Rb gene and any mismatches in the resulting heteroduplexes are identified.

Also, the isolated normal human retinoblastoma gene can be used to produce the normal Rb gene product for protein therapy of individuals determined to have a defective Rb allele.

In another aspect, the invention features a method of detecting the presence, in a tumor sample, of a protein the absence of which is associated with a distinct set of neoplasms. The method comprises producing an antibody to the Rb protein, contacting the antibody with the tumor sample, and detecting immune complexes as an indication of the presence of the protein in the tumor sample. If a tumor lacks the Rb gene product, no immune complexes will be found, and one may conclude that the tumor was the result of mutant Rb alleles. This limits the pathologic diagnosis to those tumors known to be caused by mutant Rb alleles, such as retinoblastoma, osteosarcoma, and some undifferentiated tumors of unknown cellular origin. A more exact categorization of pathologic diagnosis of human tumors will result.

Description of the Preferred Embodiments

The drawings first will be briefly described.

Drawings

Fig. 1 is a pictorial representation of the autoradiogram from a Northern blot probed with p7H30.7R;

Fig. 2 is a diagrammatic representation of the restriction map of the insert in the clone p4.7R;

Fig 3 is a pictorial representation of the autoradiogram from a Northern blot probed with p4.7R;

Fig. 4 is a diagrammatic representation of the vectors p2AR3.8 and p2AR0.9 of the invention;

Fig. 5 is a diagrammatic representation of the mismatch detection technique;

Fig. 6 is a diagrammatic representation of an example denaturing gel used in mismatch detection.

Fig. 7 is the sequence of the normal Rb gene, with flanking regions.

Isolation of the Normal Rb Gene

The genetic locus involved in causing retinoblastoma has been assigned to the q14 band of human chromosome 13 (Sparkes et al., Science 208:1042 (1980). A cDNA clone, p4.7R, from this region of DNA has been shown to carry Rb gene sequences. This clone was obtained by the following general techniques.

Isolation of cDNA Clone p4.7R

The human DNA probe pH3-8, isolated from a human chromosome 13 lambda phage library (Lalande et al., 1984, Cancer Genet. Cytogenet. 13:283), was used in a chromosome walking technique to isolate and map 30 kb of genomic DNA surrounding the H3-8 sequence. One fragment generated by this technique, named p7H30.7R, was found to recognize a DNA sequence in the mouse genome as well as within human chromosome 13 (Dryja et al., 1986, Proc. Nat. Acad. Sci. USA in press). The homology of p7H30.7R to both human and mouse DNA suggested that p7H30.7R contained coding sequences of a structural gene.

To test this possibility, p7H30.7R was radiolabeled and used to probe a Northern blot of RNA isolated from three retinoblastoma tumors (#42, #30, and #31) and an adenovirus 12-transformed human embryonic retinal cell line (Ret) (Vaessen et al., 1986, EMBO Journal 5:335). The p7H30.7R probe hybridized to an RNA transcript of approximately 4.7 kb from the retinal cell line, but did not hybridize to any RNA transcripts from the three tumor samples (Fig. 1).

Subsequently, RNA isolated from the adenovirus-transformed retinal cell line was used to construct a cDNA library. This library was screened with the labeled p7H30.7R probe. Several cDNA clones were isolated which had similar restriction maps. The longest of these, p4.7R, contained 4.7 kb of genomic DNA. The physical map of p4.7R is shown in Fig. 2. Characterization of p4.7R

The p4.7R clone was used to screen RNA transcripts isolated from retinoblastomas (#42, #30, #41, #31), an osteosarcoma (#16), and the adenovirus-transformed retinal cells (Ret). As shown in Fig. 3, the p4.7R probe detected, in a Northern blot analysis of isolated RNA's, a transcript in the transformed retinal cells which is not present in the four retinoblastoma and one osteosarcoma cell samples. The bands at ~ 2.0 kb were detected by reprobing the Northern blot, after washing, with a probe that detects rat tubulin (to demonstrate the presence of RNA in the blot).

The p4.7R clone also was used to screen genomic DNA. DNA was isolated from a set of tumors from 50 unrelated individuals, consisting of 40 retinoblastomas, 8 osteosarcomas, and 2 undifferentiated tumors of unknown cellular origin arising in patients with hereditary retinoblastoma. The isolated samples of DNA were digested with HindIII and analyzed by Southern blot hybridization using radiolabeled p4.7R as the probe. This analysis revealed three types of deviant patterns of the genomic DNA restriction fragments: totally absent fragments, representing apparent homozygous deletions; under-represented fragments, representing apparent heterozygous deletions; and fragments of altered size, reflecting either partial deletion or an alteration of a restriction site. At least 30% of the tumor DNA's exhibited one of these abnormalities. In comparison, Southern blot analysis of leukocyte DNA from 18 normal individuals showed a uniform pattern of restriction fragments.

The above results indicate that p4.7R detects the Rb gene. The deletion pattern in one osteosarcoma DNA sample provided particularly good evidence that p4.7R detects the Rb gene. This DNA sample is homozygous for a deletion that maps entirely within the p4.7R region. It is highly unlikely that the osteosarcoma phenotype arose due to a mutation independent of this deletion. Since the deletion is limited to the p4.7R region, this region must contain the Rb gene which, when mutated, produces non-functional Rb-encoded protein. The absence of functional Rb protein allows the neoplastic phenotype to develop.

Use

The p4.7R sequences can be used, according to the invention, to screen individuals for the presence of a mutated
allele of the Rb gene. This screening procedure will allow individuals having a risk of developing retinoblastoma--because of family history or a previous incidence of retinoblastoma in one eye--to determine the need for routine testing by the current ocular examination procedure. Only if the screening procedure determines that the individual possesses a mutant Rb allele will the examination procedure need to be conducted on a regular basis. Those with two normal Rb alleles can discontinue examination, as the risk of developing retinoblastoma in an individual with two normal copies of the Rb gene is approximately 1 in

20,000, or 0.005%, compared to a risk of 80%-90% if an individual has an Rb allele containing a mutation sufficient to inactivate the allele. Thus, a substantial percentage of individuals who are currently examined regularly are not actually at a greater risk than the general population: neither a family history of nor a previous incidence of retinoblastoma is conclusive evidence that an individual has the genetic predisposition to the disease. Therefore, such individuals, actually carrying two normal copies of the Rb gene, have been repeatedly undergoing the expensive and traumatic ocular examination procedure needlessly.

The screening procedure according to the invention preferably can be of two major types: (1) testing an individual's DNA for deletions in the Rb locus large enough to interfere with hybridization to an Rb probe, and (2) testing an individual's DNA for small deletions or point mutations in the Rb locus.

In addition to screening, the invention has the potential to provide protein therapy for those individuals determined to contain a mutant Rb allele and who therefore are at risk of developing retinoblastoma.

An additional use of the invention, as mentioned above, is in immunodiagnosis to determine, for example, whether a certain tumor is the result of an Rb gene abnormality. Since osteosarcomas and certain undifferentiated tumors can result from detectable lesions in the Rb gene, the immunodiagnosis can be used to aid in the diagnosis of such tumors.

Illustrative examples are given below.

Example 1: Southern Blot Analysis

To detect large deletions in the Rb locus, a Southern Blot analysis is carried out on DNA obtained from an individual to be tested. The DNA is obtained from peripheral leucocytes or, if the patient has had a tumor in one eye, from the tumor. To examine leucocyte DNA, a 10 ml blood sample is obtained from the individual, and the genomic DNA is isolated from the leucocytes in the sample, according to standard techniques. This DNA is digested with a restriction endonuclease, e.g., HindIII, run on an agarose electrophoresis gel, and transferred to a nitrocellulose filter by blotting. The DNA on the filter is then probed with radiolabeled p2AR3.8 and, separately, p2AR0.9, containing subfragments from p4.7R obtained by EcoR1 digestion (Fig. 4); it is preferred to use two or more subfragments separately rather the entire p4.7R insert, in order to better define the location of any abnormalities detected. Autoradiograms of the probed filter give a restriction map of the Rb locus in the somatic or tumor DNA of the tested individual.

This restriction map then is compared with a control restriction map, determined by using the same restriction enzyme digestion and probe. A suitable control can be DNA obtained from the adenovirus-transformed retinal cell line or leucocyte DNA from a set of normal individuals. If the tested individual has an Rb allele containing a significantly large deletion, the above restriction map of his DNA, compared with the control, will contain an additional band or bands, and/or a band or bands that have lost 50% of their intensity, caused by a change in the size, or total elimination, of one or more restriction fragments by the deletion in one allele at the Rb locus.

Thus, this screening procedure by Southern analysis will detect the existence of non-functional Rb alleles which have large deletions. If this analysis indicates that the tested DNA from an individual has a restriction map different from the control map, there is a great probability that the individual contains a non-functional, mutant Rb allele. The individual must be monitored closely thereafter for the development of retinoblastoma.

If the test restriction map appears identical to the control, a different screening procedure can be performed on the individual's DNA to determine if the individual contains an Rb allele having a small deletion or point mutation, which is sufficient to inactivate the allele but not to prevent hybridization with a probe. This screening procedure is described in the following example.

Example 2: Rb Locus Fine Structural Analysis

To examine an individual's DNA for small deletions or point mutations in the Rb locus, both homologs of the Rb gene from the individual preferably are cloned. The cloned alleles then can be tested for the presence of sequence differences from the normal allele, represented by p4.7R, by one of the following two methods: (1) the nucleotide sequence of both the cloned alleles and p4.7R are determined and then compared, or (2) RNA transcripts from p4.7R are hybridized to single stranded whole genomic DNA from an individual to be tested, and the resulting heteroduplex is treated with RNase A and run on a denaturing gel to detect the location of any mismatches. In more detail, these methods are carried out as follows:

(1) Cloning Rb alleles

The alleles of the Rb gene in an individual to be tested are cloned using conventional techniques. A common method, for example, employs the bacteriophage vector EMBL3 (Frischauf et al., 1983, J. Mol. Biol. 170:827). A 10 ml blood sample is obtained from the individual, and the genomic

DNA is isolated from the cells in this sample. This DNA is partially digested with MboI to an average fragment size of approximately 20 kb. Fragments in the range from 18-21 kb are isolated. The resulting MboI-ended fragments are ligated into the EMBL3 vector DNA which has been completely digested with BamHI, treated with alkaline phosphatase, and heated to 68°C for 10 minutes to disrupt the cohesive ends. This ligation mix is used in an in vitro lambda packaging reaction, and the packaged phage are amplified by growing a plate stock. [This cloning technique is described generally in Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Publications, pp 256-293 (1982).]

Approximately 5 x 10^5 pfu from this plate stock are used to infect 3 ml of E. coli cells at ~ 1.5 x 10^9 cells/ml in 0.01M $MgSO_4$, and the infection mix is incubated at 37°C for 20 minutes. 65 ml melted top agar at 47°C is added, and the mixture is plated onto ten 150 mm plates containing freshly poured and dry bottom agar. The agar plates are incubated until the plagues reach a diamter of ~ 1.5 mm and are just beginning to contact one another (approximately 10-12 hours).

Duplicate circular nitrocellulose filters (Millipore HAWP) are placed gently on the surface of each agar plate to bind the bacteriophage DNA. The filters are carefully removed after 1 minute, placed into a denuturing solution (1.5m NaCl, 0.5M NaOH) for 30 seconds, neutralized for 5 minutes (1.5M NaCl, 0.5M Tris-Cl pH 8.0), and dried under vacuum at 80°C for 2 hours.

These nitrocellulose filters then are probed with radiolabeled p4.7R by hybridization and autoradiography. Plaques which show hybridization to the p4.7R probe are plague-purified and rescreened according to the above procedure. Positive plagues from the rescreening are isolated and used to prepare DNA putatively containing Rb alleles from the individual.

The MboI genomic inserts in these isolated EMBL3 vector DNA samples are tested for the location of the sequences homologous to p4.7R by Southern analysis. DNA samples containing the entire Rb gene region are selected, and the appropriate restriction fragments containing the Rb gene from these samples are subcloned into a suitable vector, such as pUC9. These subclones thus contain copies of one or both Rb alleles from the DNA of the individual to be tested. To determine if both alleles are represented, the initial phage isolates are tested for the existence of restriction polymorphism. These subcloned alleles are then examined for differences from p4.7R by one of the following techniques.

(2) Sequence Comparison

First, the nucleotide sequence of the normal Rb gene in p4.7R is determined by subcloning restriction fragments of ~ 500 bp from p4.7R into an M13mp8 phage vector and sequencing these sublones by the dideoxy technique (Sanger et al., 1977, Proc. Nat. Acad. Sci USA 74:5463). A composite sequence of the Rb gene then can be assembled from these individual subclone sequences. This sequence is given in Fig. 7 which also shows flanking regions.

The isolated Rb gene alleles are sequenced according to the following procedure. Restriction fragments (~ 2kb) of the allele are subcloned into the M13mp8 vector, and short stretches (~500 bp) are sequenced individually using small restriction fragments isolated from p4.7R as the primers in the dideoxy sequencing reactions. The composite nucleotide sequence of the isolated allele then can be constructed from these individually-primed sequences. This sequence is compared directly with the sequence of the normal Rb gene, determined from p4.7R, to determine if any deletions or point mutations exist in the isolated allele.

(3) Ribonuclease Cleavage of Mismatches

An alternative method of comparing the allelic DNA with the normal Rb gene employs RNase A to detect the existence of differences between the p4.7R sequence and the allele sequence. This comparison is performed in steps with small (~ 500 bp) restriction fragments of p4.7R as the probe. First, p4.7R is digested with a restriction enzyme(s) that cuts the Rb gene sequence into fragments of approximately 500bp. These fragements are isolated on an electrophoresis gel and cloned individually, in both orientations, into an SP6 vector, such as pSP64 or pSP65 (Melton et al., 1984, Nucleic Acids Res. 12:7035). The SP6-based plasmids containing inserts of p4.7R fragments are transcribed in vitro using the SP6 transcription system, well known in the art, in the presence of [α-^{32}P]GTP, generating radiolabeled RNA transcripts of both strands of the cDNA of the Rb gene.

Individually, these RNA transcripts are used to form heteroduplexes with the allelic DNA, as follows. 50 ng of the allele subclone is digested with a restriction enzyme that cuts outside of the region covered by the RNA transcript probe to be used. This digested DNA is mixed with the radiolabeled RNA probe in 30 µl of hybrization buffer (80% formamide, 40 mM Pipes pH6.4, 0.4M NaCl, and 1mM EDTA) and the mixture is treated at 90°C for 10 minutes to denature the DNA. The mixture then is cooled slowly to 45°C and the RNA is allowed to anneal to the single-stranded DNA at 45°C for half

an hour.

The RNA:DNA heteroduplexes next are treated with 350 µl of an RNase A solution (Sigma) (40 µg/ml in 10mM Tris-HCl pH7.5, 1mM EDTA, 0.2M NaCl, and 0.1M LiCl). The mixture is vortexed and incubated at 25°C for 30 minutes. The RNase A reaction is stopped by adding 10µl of proteinase K (10mg/ml) (Boehringer Mannheim) followed by incubation at 37°C for 20 minutes. Extraction with phenol-chloroform and ethanol precipitation of the aqueous layer yields a nucleic acid sample free from protein contamination. The precipitated sample is resuspended in 5µl and analyzed by denaturing polyacrylamide gel electrophoresis (4% polyacrylamide, 7M urea) (Fig. 5).

Mismatches that occur in the RNA:DNA heteroduplex, due to sequence differences between the p4.7R fragment and the Rb allele subclone from the individual, result in cleavage in the RNA strand by the RNase A treatment. Such mismatches can be the result of point mutations or small deletions in the individual's Rb allele. Cleavage of the RNA strand yields two or more small RNA fragments, which run faster on the denaturing gel than the RNA probe itself, as shown in Fig. 6.

In the above RNAse A technique, radiolabeled Rb gene RNA is hybridized to single strands of an individual's Rb allele which has been cloned into a vector. The RNase A technique is advantageous, however, because it also can be used without having to clone the Rb alleles. Preferably, genomic DNA is isolated from blood cells of the individual to be tested, and this genomic DNA is hybridized directly with the radiolabeled Rb RNA probes to determine sequence differences from the normal Rb gene, as follows. 5 µg of isolated, total genomic DNA is resuspended with the labeled RNA probe in 30 µl of hybridization buffer (80% formamide, 40mM Pipes pH6.4, 0.HM NaCl, and 1mM EDTA), and this hybridization mix is treated at 90°C for 10 minutes to denature the DNA. The mixture then is cooled slowly to 45°C and incubated at this temperature for 10 hours to allow hybridization of the RNA probe to the single-stranded DNA copies of the Rb allele. After hybridization, the RNase A treatment and electophoresis are performed as above. Mismatches in the heteroduplexes between the RNA probe and the genomic copies of the individual's Rb alleles are readily detected.

Example 3: Protein Therapy

Another use for the cloned cDNA of the normal Rb gene, as represented by p4.7R, is to produce the Rb protein for treatment of individuals determined to carry a defective allele of the Rb gene. To prevent the formation of retinoblastoma in these individuals, the Rb gene product is administered therapeutically to these individuals. The Rb protein is produced by cloning the Rb cDNA from p4.7R into an appropriate mammalian expression vector, expressing the Rb protein from this vector in an in vivo expression system, and isolating the Rb protein from the medium or cells of the expression system.

General in vitro expression vectors and systems are well known in the art.

Example 4: Immunodiagnosis

The Rb protein, produced as described above, is injected into a rabbit to produce anti-Rb antibody, which then is labeled, e.g., radioactively, fluorescently, or with an enzyme such as alkaline phosphatase. The labeled antibody is used to determine whether human tumors are of defective Rb gene origin. This can be carried out using any conventional technique. For example, the tumor sample can be liquified and tested against the labeled antibody using a conventional ELISA format. Alternatively, a tumor section can be fixed and reacted with labeled antibody, and any immune complexes then can be detected by autoradiography or fluorescence microscopy, depending on the type of label on the antibody. Tumors lacking an antigen reactive with the antibody to the Rb gene product are due to mutations of the retinoblastoma gene. Since the tumors known to be caused by a mutant Rb gene are few (including retinoblastoma and osteosarcoma), the differential diagnosis of tumors deficient in the Rb gene product is greatly limited by such a test.

Deposits

The plasmids p2AR3.8 and p2AR0.9 were deposited on July 17, 1987 with the American Type Culture Collection, Rockville, Maryland, and assigned ATCC accession numbers 40,241 and 40,242, respectively.

Other embodiments are within the following claims.

Claims

1. A method of detecting the presence, in a tumour sample, of a protein the absence of which is associated with a neoplasm, the method comprising producing an antibody to the protein, contacting the antibody with the tumour sample, and detecting immune complexes as an indication of the presence in the tumour sample of the protein.

2. A method according to claim 1 wherein the neoplasm is caused by a mutant Rb allele.

3. A method according to claim 2 wherein the mutant allele is a result of retinoblastoma or osteosarcoma.

4. A method according to any preceding claim wherein the antibody is labelled.

5. An immunodiagnostic kit comprising a labelled antibody specific for a protein associated with a neoplasm, the absence of which protein indicates the presence of a tumour.

Claims for the following Contracting States : AT, ES

1. A method of detecting the presence, in a tumour sample, of a protein the absence of which is associated with a neoplasm, the method comprising producing an antibody to the protein, contacting the antibody with the tumour sample, and detecting immune complexes as an indication of the presence in the tumour sample of the protein.

2. A method according to claim 1 wherein the neoplasm is caused by a mutant Rb allele.

3. A method according to claim 2 wherein the mutant allele is a result of retinoblastoma or osteosarcoma.

4. A method according to any preceding claim wherein the antibody is labelled.

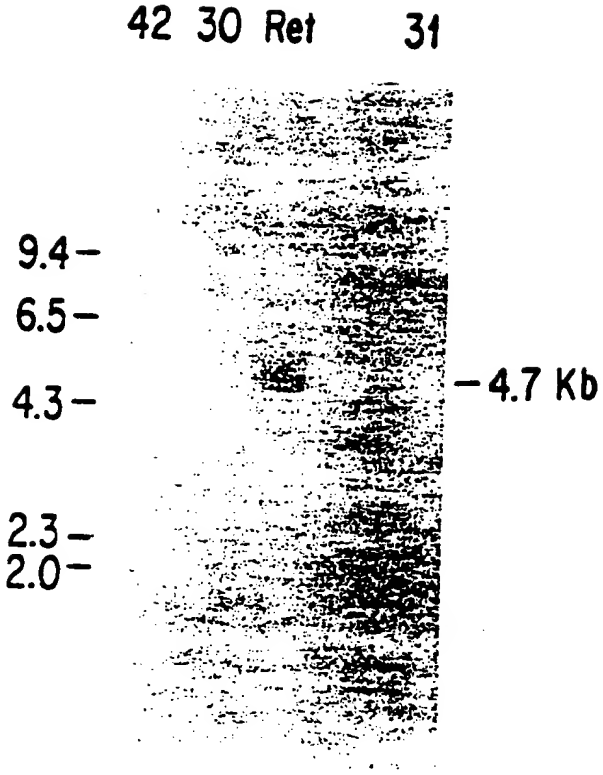

FIG. 1

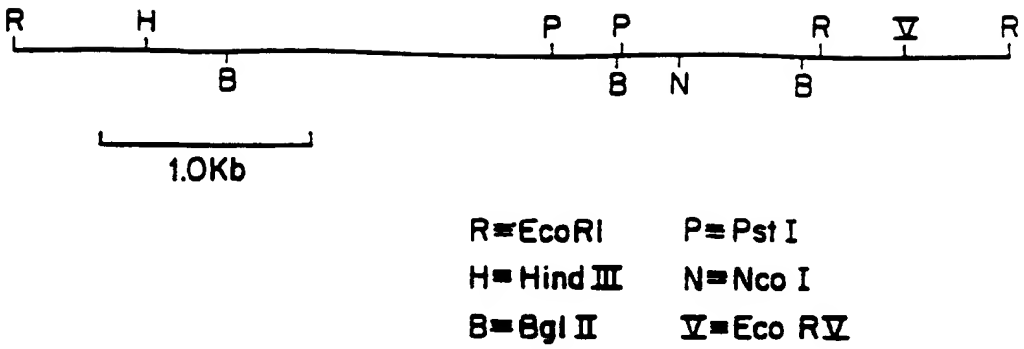
R
H
P
P
R
V
R
B
B
N
B
1.0Kb
R=EcoRI
P=Pst I
H=Hind III
N=Nco I
B=Bgl II
V=Eco RV

FIG. 2

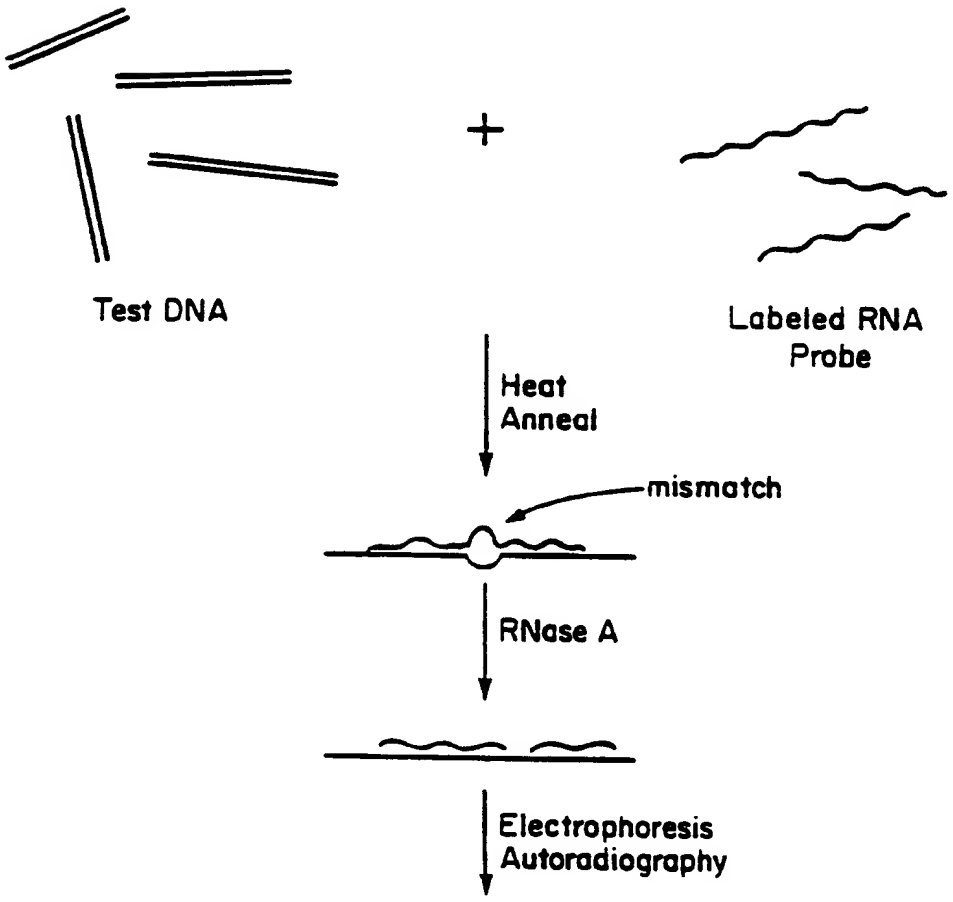
Test DNA
Labeled RNA Probe
Heat Anneal
mismatch
RNase A
Electrophoresis Autoradiography

FIG. 5

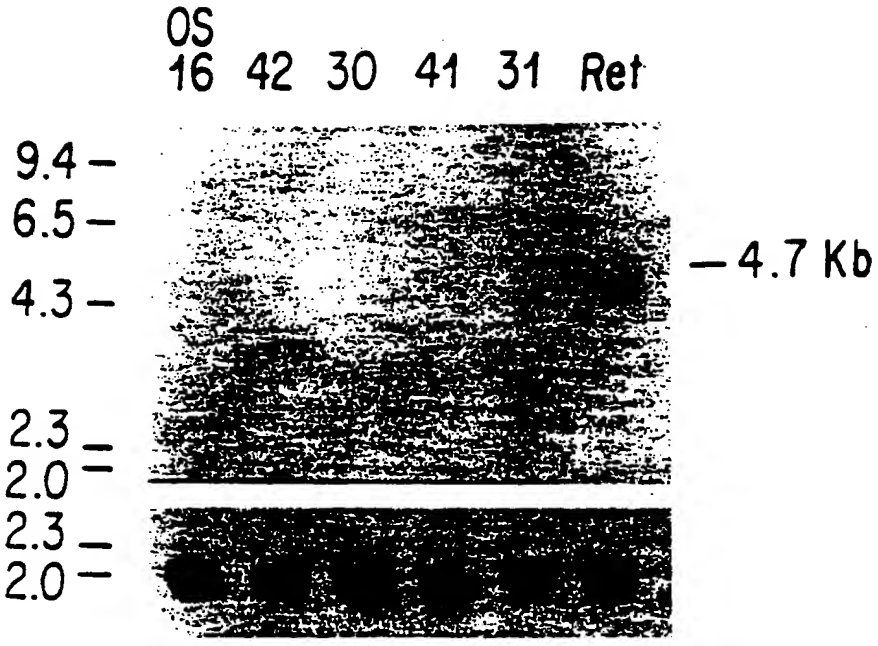

FIG. 3

FIG. 4

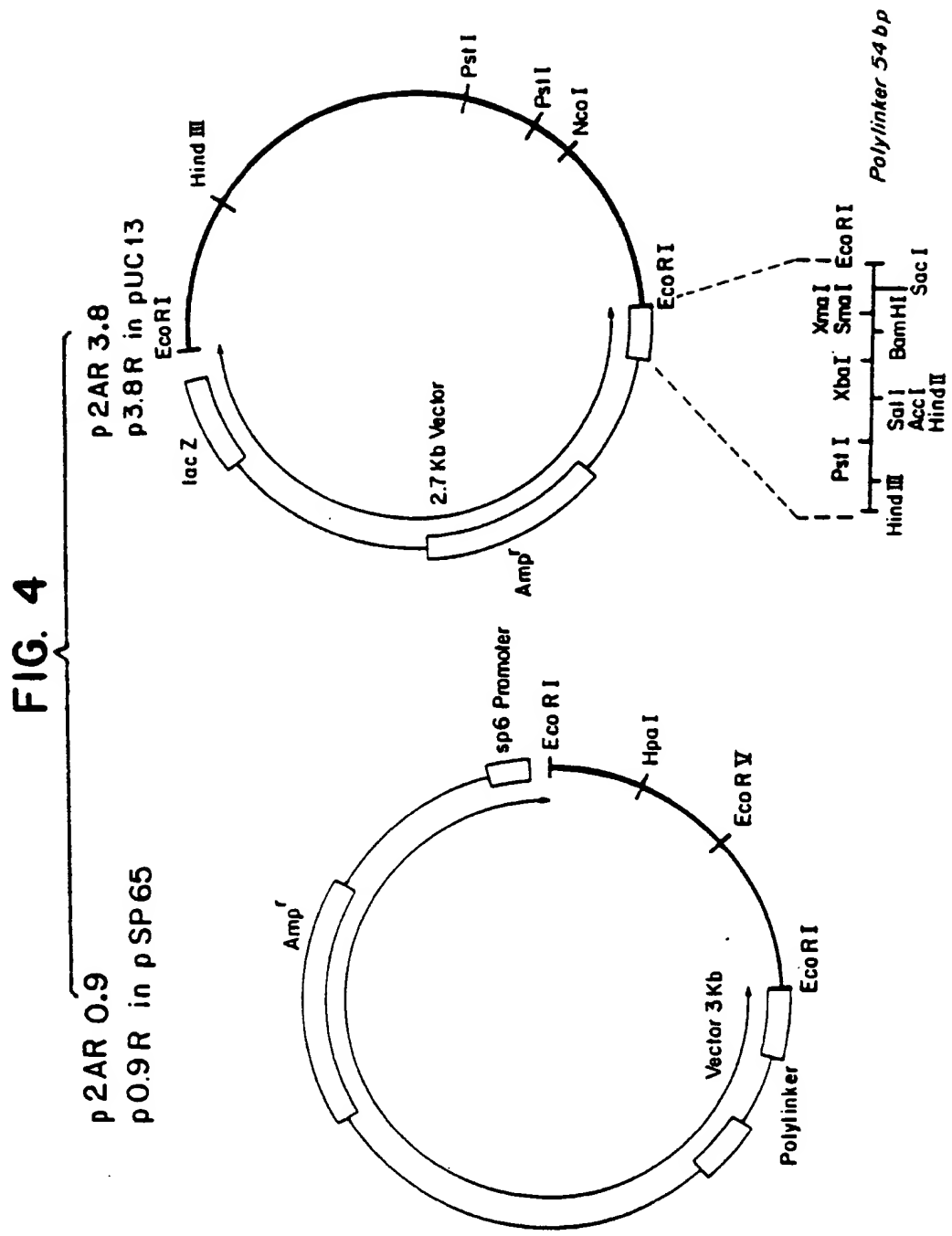
p2AR 0.9
p0.9R in pSP65
p2AR 3.8
p3.8R in pUC13
Amp^r
sp6 Promoter
Eco R I
Hpa I
Eco R V
Eco R I
Polylinker
Vector 3Kb
Eco R I
Hind III
Pst I
Pst I
Nco I
lac Z
2.7 Kb Vector
Amp^r
Eco R I
Polylinker 54 bp
Xma I
Sma I
Eco R I
Sac I
Pst I
Xba I
Bam HI
Sal I
Acc I
Hind II
Hind III

FIG. 6

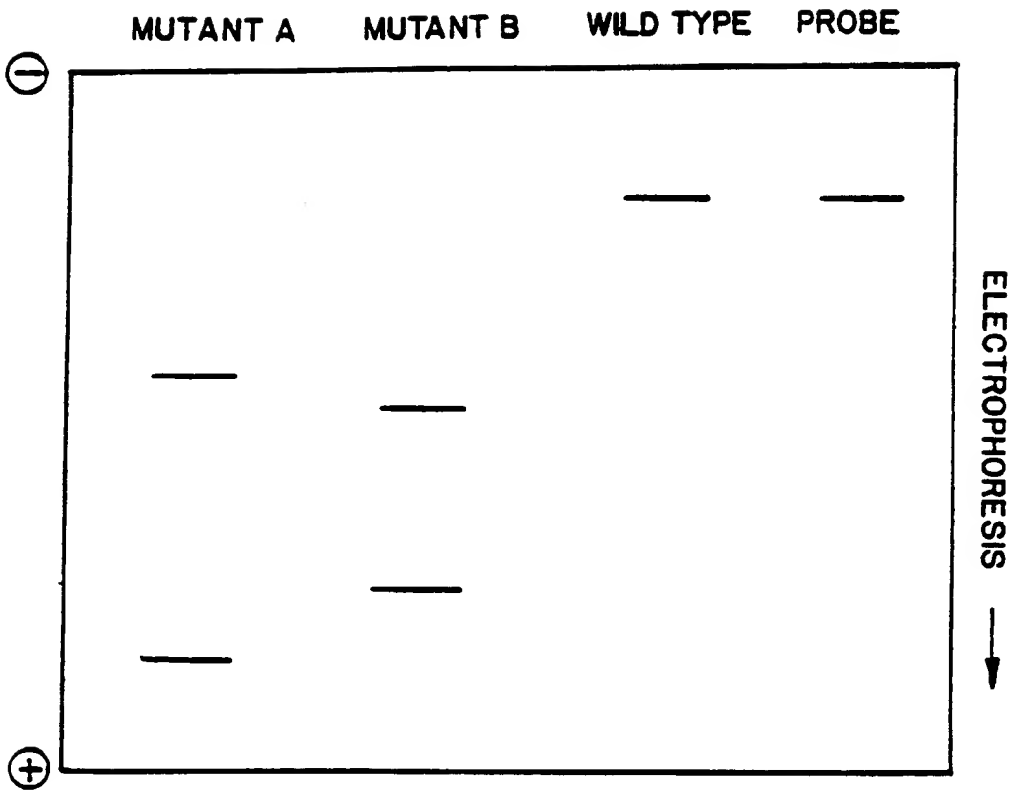

FIG. 7-1

```
                10                  20                  30                  40                  50                  60
GTC ATG CCG CCC AAA ACC CCC CGA AAA ACG GCC GCC ACC GCC GCC GCT GCC GCC GCG GAA
CCC CCG GCA CCG GCC GCC GCC CCC TCC TGA GGA GGA CCC AGA GCA GGA CAG CGG CCC GGA
               130                 140                 150                 160                 170                 180
GAC CTG CCT CTC GTC AGG CTT GAG TTT GAA GAA ACA GAA GAA CCT GAT TTT ACT GCA TTA
               190                 200                 210                 220                 230                 240
TGT CAG AAA TTA AAG ATA CCA GAT CAT GTC AGA GAG AGA GCT TGG TTA ACT TGG GAG AAA
               250                 260                 270                 280                 290                 300
GTT TCA TCT GTG GAT GGA GTA TTG GGA GGT TAT ATT CAA AAG AAA AAG GAA CTG TGG GGA
               310                 320                 330                 340                 350                 360
ATC TGT ATC TTT ATT GCA GCA GTT GAC CTA GAT GAG ATG TCG TTC ACT TTT ACT GAG CTA
                                                        M   S   F   T   F   T   E   L

               370                 380                 390                 400                 410                 420
CAG AAA AAC ATA GAA ATC AGT GTC CAT AAA TTC TTT AAC TTA CTA AAA GAA ATT GAT ACC
 Q   K   N   I   E   I   S   V   H   K   F   F   N   L   L   K   E   I   D   T

               430                 440                 450                 460                 470                 480
AGT ACC AAA GTT GAT AAT GCT ATG TCA AGA CTG TTG AAG AAG TAT GAT GTA TTG TTT GCA
 S   T   K   V   D   N   A   M   S   R   L   L   K   K   Y   D   V   L   F   A

               490                 500                 510                 520                 530                 540
CTC TTC AGC AAA TTG GAA AGG ACA TGT GAA CTT ATA TAT TTG ACA CAA CCC AGC AGT TCG
 L   F   S   K   L   E   R   T   C   E   L   I   Y   L   T   Q   P   S   S   S

               550                 560                 570                 580                 590                 600
ATA TCT ACT GAA ATA AAT TCT GCA TTG GTG CTA AAA GTT TCT TGG ATC ACA TTT TTA TTA
 I   S   T   E   I   N   S   A   L   V   L   K   V   S   W   I   T   F   L   L

               610                 620                 630                 640                 650                 660
GCT AAA GGG GAA GTA TTA CAA ATG GAA GAT GAT CTG GTG ATT TCA TTT CAG TTA ATG CTA
 A   K   G   E   V   L   Q   M   E   D   D   L   V   I   S   F   Q   L   M   L

               670                 680                 690                 700                 710                 720
TGT GTC CTT GAC TAT TTT ATT AAA CTC TCA CCT CCC ATG TTG CTC AAA GAA CCA TAT AAA
 C   V   L   D   Y   F   I   K   L   S   P   P   M   L   L   K   E   P   Y   K

               730                 740                 750                 760                 770                 780
ACA GCT GTT ATA CCC ATT AAT GGT TCA CCT CGA ACA CCC AGG CGA GGT CAG AAC AGG AGT
 T   A   V   I   P   I   N   G   S   P   R   T   P   R   R   G   Q   N   R   S
```

FIG. 7-2

```
          790        800        810        820        830        840
           *          *          *          *          *          *
GCA CGG ATA GCA AAA CAA CTA GAA AAT GAT ACA AGA ATT ATT GAA GTT CTC TGT AAA GAA
 A   R   I   A   K   Q   L   E   N   D   T   R   I   I   E   V   L   C   K   E

          850        860        870        880        890        900
           *          *          *          *          *          *
CAT GAA TGT AAT ATA GAT GAG GTG AAA AAT GTT TAT TTC AAA AAT TTT ATA CCT TTT ATG
 H   E   C   N   I   D   E   V   K   N   V   Y   F   K   N   F   I   P   F   M

          910        920        930        940        950        960
           *          *          *          *          *          *
AAT TCT CTT GGA CTT GTA ACA TCT AAT GGA CTT CCA GAG GTT GAA AAT CTT TCT AAA CGA
 N   S   L   G   L   V   T   S   N   G   L   P   E   V   E   N   L   S   K   R

          970        980        990       1000       1010       1020
           *          *          *          *          *          *
TAC GAA GAA ATT TAT CTT AAA AAT AAA GAT CTA GAT GCA AGA TTA TTT TTG GAT CAT GAT
 Y   E   E   I   Y   L   K   N   K   D   L   D   A   R   L   F   L   D   H   D

         1030       1040       1050       1060       1070       1080
           *          *          *          *          *          *
AAA ACT CTT CAG ACT GAT TCT ATA GAC AGT TTT GAA ACA CAG AGA ACA CCA CGA AAA AGT
 K   T   L   Q   T   D   S   I   D   S   F   E   T   Q   R   T   P   R   K   S

         1090       1100       1110       1120       1130       1140
           *          *          *          *          *          *
AAC CTT GAT GAA GAG GTG AAT GTA ATT CCT CCA CAC ACT CCA GTT AGG ACT GTT ATG AAC
 N   L   D   E   E   V   N   V   I   P   P   H   T   P   V   R   T   V   M   N

         1150       1160       1170       1180       1190       1200
           *          *          *          *          *          *
ACT ATC CAA CAA TTA ATG ATG ATT TTA AAT TCA GCA AGT GAT CAA CCT TCA GAA AAT CTG
 T   I   Q   Q   L   M   M   I   L   N   S   A   S   D   Q   P   S   E   N   L

         1210       1220       1230       1240       1250       1260
           *          *          *          *          *          *
ATT TCC TAT TTT AAC AAC TGC ACA GTG AAT CCA AAA GAA AGT ATA CTG AAA AGA GTG AAG
 I   S   Y   F   N   N   C   T   V   N   P   K   E   S   I   L   K   R   V   K

         1270       1280       1290       1300       1310       1320
           *          *          *          *          *          *
GAT ATA GGA TAC ATC TTT AAA GAG AAA TTT GCT AAA GCT GTG GGA CAG GGT TGT GTC GAA
 D   I   G   Y   I   F   K   E   K   F   A   K   A   V   G   Q   G   C   V   E

         1330       1340       1350       1360       1370       1380
           *          *          *          *          *          *
ATT GGA TCA CAG CGA TAC AAA CTT GGA GTT CGC TTG TAT TAC CGA GTA ATG GAA TCC ATG
 I   G   S   Q   R   Y   K   L   G   V   R   L   Y   Y   R   V   M   E   S   M

         1390       1400       1410       1420       1430       1440
           *          *          *          *          *          *
CTT AAA TCA GAA GAA GAA CGA TTA TCC ATT CAA AAT TTT AGC AAA CTT CTG AAT GAC AAC
 L   K   S   E   E   E   R   L   S   I   Q   N   F   S   K   L   L   N   D   N
```

FIG. 7-3

```
          1450        1460        1470        1480        1490        1500
           *           *           *           *           *           *
ATT TTT CAT ATG TCT TTA TTG GCG TGC GCT CTT GAG GTT GTA ATG GCC ACA TAT AGC AGA
 I   F   H   M   S   L   L   A   C   A   L   E   V   V   M   A   T   Y   S   R

          1510        1520        1530        1540        1550        1560
           *           *           *           *           *           *
AGT ACA TCT CAG AAT CTT GAT TCT GGA ACA GAT TTG TCT TTC CCA TGG ATT CTG AAT GTG
 S   T   S   Q   N   L   D   S   G   T   D   L   S   F   P   W   I   L   N   V

          1570        1580        1590        1600        1610        1620
           *           *           *           *           *           *
CTT AAT TTA AAA GCC TTT GAT TTT TAC AAA GTG ATC GAA AGT TTT ATC AAA GCA GAA GGC
 L   N   L   K   A   F   D   F   Y   K   V   I   E   S   F   I   K   A   E   G

          1630        1640        1650        1660        1670        1680
           *           *           *           *           *           *
AAC TTG ACA AGA GAA ATG ATA AAA CAT TTA GAA CGA TGT GAA CAT CGA ATC ATG GAA TCC
 N   L   T   R   E   M   I   K   H   L   E   R   C   E   H   R   I   M   E   S

          1690        1700        1710        1720        1730        1740
           *           *           *           *           *           *
CTT GCA TGG CTC TCA GAT TCA CCT TTA TTT GAT CTT ATT AAA CAA TCA AAG GAC CGA GAA
 L   A   W   L   S   D   S   P   L   F   D   L   I   K   Q   S   K   D   R   E

          1750        1760        1770        1780        1790        1800
           *           *           *           *           *           *
GGA CCA ACT GAT CAC CTT GAA TCT GCT TGT CCT CTT AAT CTT CCT CTC CAG AAT AAT CAC
 G   P   T   D   H   L   E   S   A   C   P   L   N   L   P   L   Q   N   N   H

          1810        1820        1830        1840        1850        1860
           *           *           *           *           *           *
ACT GCA GCA GAT ATG TAT CTT TCT CCT GTA AGA TCT CCA AAG AAA AAA GGT TCA ACT ACG
 T   A   A   D   M   Y   L   S   P   V   R   S   P   K   K   K   G   S   T   T

          1870        1880        1890        1900        1910        1920
           *           *           *           *           *           *
CGT GTA AAT TCT ACT GCA AAT GCA GAG ACA CAA GCA ACC TCA GCC TTC CAG ACC CAG AAG
 R   V   N   S   T   A   N   A   E   T   Q   A   T   S   A   F   Q   T   Q   K

          1930        1940        1950        1960        1970        1980
           *           *           *           *           *           *
CCA TTG AAA TCT ACC TCT CTT TCA CTG TTT TAT AAA AAA GTG TAT CGG CTA GCC TAT CTC
 P   L   K   S   T   S   L   S   L   F   Y   K   K   V   Y   R   L   A   Y   L

          1990        2000        2010        2020        2030        2040
           *           *           *           *           *           *
CGG CTA AAT ACA CTT TGT GAA CGC CTT CTG TCT GAG CAC CCA GAA TTA GAA CAT ATC ATC
 R   L   N   T   L   C   E   R   L   L   S   E   H   P   E   L   E   H   I   I

          2050        2060        2070        2080        2090        2100
           *           *           *           *           *           *
TGG ACC CTT TTC CAG CAC ACC CTG CAG AAT GAG TAT GAA CTC ATG AGA GAC AGG CAT TTG
 W   T   L   F   Q   H   T   L   Q   N   E   Y   E   L   M   R   D   R   H   L

          2110        2120        2130        2140        2150        2160
           *           *           *           *           *           *
GAC CAA ATT ATG ATG TGT TCC ATG TAT GGC ATA TGC AAA GTG AAG AAT ATA GAC CTT AAA
 D   Q   I   M   M   C   S   M   Y   G   I   C   K   V   K   N   I   D   L   K

          2170        2180        2190        2200        2210        2220
           *           *           *           *           *           *
TTC AAA ATC ATT GTA ACA GCA TAC AAG GAT CTT CCT CAT GCT GTT CAG GAG ACA TTC AAA
 F   K   I   I   V   T   A   Y   K   D   L   P   H   A   V   Q   E   T   F   K

          2230        2240        2250        2260        2270        2280
           *           *           *           *           *           *
CGT GTT TTG ATC AAA GAA GAG GAG TAT GAT TCT ATT ATA GTA TTC TAT AAC TCG GTC TTC
 R   V   L   I   K   E   E   E   Y   D   S   I   I   V   F   Y   N   S   V   F
```

FIG. 7-4

```
          2290          2300          2310          2320          2330          2340
            *             *             *             *             *             *
ATG CAG AGA CTG AAA ACA AAT ATT TTG CAG TAT GCT TCC ACC AGG CCC CCT ACC TTG TCA
 M   Q   R   L   K   T   N   I   L   Q   Y   A   S   T   R   P   P   T   L   S

          2350          2360          2370          2380          2390          2400
            *             *             *             *             *             *
CCA ATA CCT CAC ATT CCT CGA AGC CCT TAC AAG TTT CCT AGT TCA CCC TTA CGG ATT CCT
 P   I   P   H   I   P   R   S   P   Y   K   F   P   S   S   P   L   R   I   P

          2410          2420          2430          2440          2450          2460
            *             *             *             *             *             *
GGA GGG AAC ATC TAT ATT TCA CCC CTG AAG AGT CCA TAT AAA ATT TCA GAA GGT CTG CCA
 G   G   N   I   Y   I   S   P   L   K   S   P   Y   K   I   S   E   G   L   P

          2470          2480          2490          2500          2510          2520
            *             *             *             *             *             *
ACA CCA ACA AAA ATG ACT CCA AGA TCA AGA ATC TTA GTA TCA ATT GGT GAA TCA TTC GGG
 T   P   T   K   M   T   P   R   S   R   I   L   V   S   I   G   E   S   F   G

          2530          2540          2550          2560          2570          2580
            *             *             *             *             *             *
ACT TCT GAG AAG TTC CAG AAA ATA AAT CAG ATG GTA TGT AAC AGC GAC CGT GTG CTC AAA
 T   S   E   K   F   Q   K   I   N   Q   M   V   C   N   S   D   R   V   L   K

          2590          2600          2610          2620          2630          2640
            *             *             *             *             *             *
AGA AGT GCT GAA GGA AGC AAC CCT CCT AAA CCA CTG AAA AAA CTA CGC TTT GAT ATT GAA
 R   S   A   E   G   S   N   P   P   K   P   L   K   K   L   R   F   D   I   E

          2650          2660          2670          2680          2690          2700
            *             *             *             *             *             *
GGA TCA GAT GAA GCA GAT GGA AGT AAA CAT CTC CCA GGA GAG TCC AAA TTT CAG CAG AAA
 G   S   D   E   A   D   G   S   K   H   L   P   G   E   S   K   F   Q   Q   K

          2710          2720          2730          2740          2750          2760
            *             *             *             *             *             *
CTG GCA GAA ATG ACT TCT ACT CGA ACA CGA ATG CAA AAG CAG AAA ATG AAT GAT AGC ATG
 L   A   E   M   T   S   T   R   T   R   M   Q   K   Q   K   M   N   D   S   M

          2770          2780          2790          2800          2810          2820
            *             *             *             *             *             *
GAT ACC TCA AAC AAG GAA GAG AAA TGA GGA TCT CAG GAC CTT GGT GGA CAC TGT GTA CAC
 D   T   S   N   K   E   E   K   -

          2830          2840          2850          2860          2870          2880
            *             *             *             *             *             *
CTC TGG ATT CAT TGT CTC TCA CAG ATG TGA CTG TAT AAC TTT CCC AGG TTC TGT TTA TGG

          2890          2900          2910          2920          2930          2940
            *             *             *             *             *             *
CCA CAT TTA ATA TCT TCA GCT CTT TTT GTG GAT ATA AAA TGT GCA GAT GCA ATT GTT TGG

          2950          2960          2970          2980          2990          3000
            *             *             *             *             *             *
GTG ATT CCT AAG CCA CTT GAA ATG TTA GTC ATT GTT ATT TAT ACA AGA TTG AAA ATC TTG

          3010          3020          3030          3040          3050          3060
            *             *             *             *             *             *
```

FIG. 7-5

TGT AAA TCC TGC CAT TTA AAA AGT TGT AGC AGA TTG TTT CCT CTT CCA AAG TAA AAT TGC

3070 3080 3090 3100 3110 3120
TGT GCT TTA TGG ATA GTA AGA ATG GCC CTA GAG TGG GAG TCC TGA TAA CCC AGG CCT GTC

3130 3140 3150 3160 3170 3180
TGA CTA CTT TGC CTT CTT TTG TAG CAT ATA GGT GAT GTT TGC TCT TGT TTT TAT TAA TTT

3190 3200 3210 3220 3230 3240
ATA TGT ATA TTT TTT TAA TTT AAC ATG AAC ACC CTT AGA AAA TGT GTC CTA TCT ATC TTC

3250 3260 3270 3280 3290 3300
CAA ATG CAA TTT GAT TGA CTG CCC ATT CAC CAA AAT TAT CCT GAA CTC TTC TGC AAA AAT

3310 3320 3330 3340 3350 3360
GGA TAT TAT TAG AAA TTA GAA AAA AAT TAC TAA TTT TAC ACA TTA GAT TTT ATT TTA CTA

3370 3380 3390 3400 3410 3420
TTG GAA TCT GAT ATA CTG TGT GCT TGT TTT ATA AAA TTT TGC TTT TAA TTA AAT AAA AGC

3430 3440 3450 3460 3470 3480
TGG AAG CAA AGT ATA ACC ATA TGA TAC TAT CAT ACT ACT GAA ACA GAT TTC ATA CCT CAG

3490 3500 3510 3520 3530 3540
AAT GTA AAA GAA CTT ACT GAT TAT TTT CTT CAT CCA ACT TAT GTT TTT AAA TGA GGA TTA

3550 3560 3570 3580 3590 3600
TTG ATA GTA CTC TTG GTT TTT ATA CCA TTC AGA TCA CTG AAT TTA TAA AGT ACC CAT CTA

3610 3620 3630 3640 3650 3660
GTA CTT GAA AAA GTA AAG TGT TCT GCC AGA TCT TAG GTA TAG AGG ACC CTA ACA CAG TAT

3670 3680 3690 3700 3710 3720
ATC CCA AGT GCA CTT TCT AAT GTT TCT GGG TCC TGA AGA ATT AAG ATA CAA ATT AAT TTT

3730 3740 3750 3760 3770 3780
ACT CCA TAA ACA GAC TGT TAA TTA TAG GAG CCT TAA TTT TTT TTT CAT AGA GAT TTG TCT

3790 3800 3810 3820 3830 3840
AAT TGC ATC TCA AAA TTA TTC TCC CCT CCT TAA TTT GGG AAG GTT TGT GTT TTC TCT GGA

FIG. 7-6

```
          3850        3860        3870        3880        3890        3900
           *           *           *           *           *           *
ATG GTA CAT GTC TTC CAT GTA TCT TTT GAA CTG GCA ATT GTC TAT TTA TCT TTT ATT TTT

          3910        3920        3930        3940        3950        3960
           *           *           *           *           *           *
TTA AGT CAG TAT GGT CTA ACA CTG GCA TGT TCA AAG CCA CAT TAT TTC TAG TCC AAA ATT

          3970        3980        3990        4000        4010        4020
           *           *           *           *           *           *
ACA AGT AAT CAA GGG TCA TTA TGG GTT AGG CAT TAA TGT TTC TAT CTG ATT TTG TGC AAA

          4030        4040        4050        4060        4070        4080
           *           *           *           *           *           *
AGC TTC AAA TTA AAA CAG CTG CAT TAG AAA AAG AGG CGC TTC TCC CCT CCC CTA CAC CTA

          4090        4100        4110        4120        4130        4140
           *           *           *           *           *           *
AAG GTG TAT TTA AAC TAT CTT GTG TGA TTA ACT TAT TTA GAG ATG CTG TAA CTT AAA ATA

          4150        4160        4170        4180        4190        4200
           *           *           *           *           *           *
GGG GAT ATT TAA GGT AGC TTC AGC TAG CTT TTA GGA AAA TCA CTT TGT CTA ACT CAG AAT

          4210        4220        4230        4240        4250        4260
           *           *           *           *           *           *
TAT TTT TAA AAA GAA ATC TGG TCT TGT TAG AAA ACA AAA TTT TAT TTT GTG CTC ATT TAA

          4270        4280        4290        4300        4310        4320
           *           *           *           *           *           *
GTT TCA AAC TTA CTA TTT TGA CAG TTA TTT TGA TAA CAA TGA CAC TAG AAA ACT TGA CTC

          4330        4340        4350        4360        4370        4380
           *           *           *           *           *           *
CAT TTC ATC ATT GTT TCT GCA TGA ATA TCA TAC AAA TCA GTT AGT TTT TAG GTC AAG GGC

          4390        4400        4410        4420        4430        4440
           *           *           *           *           *           *
TTA CTA TTT CTG GGT CTT TTG CTA CTA AGT TCA CAT TAG AAT TAG TGC CAG AAT TTT AGG

          4450        4460        4470        4480        4490        4500
           *           *           *           *           *           *
AAC TTC AGA GAT CGT GTA TTG AGA TTT CTT AAA TAA TGC TTC AGA TAT TAT TGC TTT ATT

          4510        4520        4530        4540        4550        4560
           *           *           *           *           *           *
GCT TTT TTG TAT TGG TTA AAA CTG TAC ATT TAA AAT TGC TAT GTT ACT ATT TTC TAC AAT

          4570        4580        4590
           *           *           *
TAA TAG TTT GTC TAT TTT AAA ATA AAT TAG TTG TTA G
```